Prediction of Clock Bias for BeiDou Satellites Using a Combination of Variational Mode Decomposition and Long Short-Term Memory Network

Wenlong Sun[1], Fengfeng Shi[1], Yuting Lin[1], Jinfeng Xu[1], Shaofeng He[2], Yongxin Lin[1], Guocheng Wang[3], Run Zhao[1] and Pei Ma[4]

1. Beijing Satellite Navigation Center, Beijing, China, 10009
2. Shenzhen Campus of Sun Yat-Sen University, Shenzhen, China, 518107
3. Innovation Academy for Precision Measurement Science and Technology, Chinese Academy of Sciences, Wuhan, China, 430077
4. Lanzhou Institute of Physics,CAST, Lanzhou, China,730000

Emails:{swl_tf@163.com, fengfeng_shi@163.com, lyt1108@163.com, ccpoppy@163.com, heshf5@mail2.sysu.edu.cn, 1029273182@qq.com, guocheng96@apm.ac.cn, zhaorun236@126.com and mig_sky@sina.com}

Abstract: The precise estimation of the satellite clock bias (SCB) holds considerable importance in ensuring accurate timekeeping, navigation, and positioning. This study introduces a novel SCB prediction approach that integrates variational mode decomposition (VMD) and long short-term memory (LSTM) network techniques, combining signal decomposition with deep learning methodologies. Initially, the raw SCB data undergoes preprocessing, followed by decomposition using the VMD method to generate multiple intrinsic mode functions (IMFs). These decomposed IMFs serve as inputs for LSTM, where several independent LSTM models are established for training and prediction purposes. Subsequently, the predicted outcomes are aggregated and reconstructed to derive the final SCB prediction. Experimental findings demonstrate notable advancements in clock bias prediction for the spaceborne hydrogen atomic clock for BDS, with prediction accuracies of 0.048 ns, 0.204 ns and 1.397 ns for 6 hours, 3 days and 15 days, respectively. These results exhibit significant enhancements compared to both the LSTM network and the Back Propagation (BP) neural network, with improvements of 56%, 84% and 83% for the aforementioned time intervals in comparison to LSTM, and enhancements of 59%, 82% and 83% relative to the BP neural network.

Keywords: Satellite Clock Bias; Prediction; VMD; LSTM; BP; Time and Frequency

1 Introduction

The atomic clock serves as a highly precise timekeeping device and is a fundamental component of satellite navigation systems. Its stability and accuracy play a critical role in ensuring precise time synchronization across various fields such as modern science, navigation systems, and communication measurements. However, the atomic clock's stability may be affected by external factors over time. The real-time clock bias data provided by the International GNSS Service (IGS) may not be sufficient for achieving real-time precision in single-point positioning of satellite navigation systems. Additionally, the subsequent release of the final precision clock bias data is typically delayed by 13 days (IGS, 2024). Therefore, the accurate forecasting of atomic clock bias is essential for achieving precise navigation and positioning, as well as for maintaining accurate timekeeping, time synchronization, and

frequency control.

In recent decades, numerous researchers have developed various models for predicting the satellite clock bias (SCB), broadly categorized into two groups: frequency domain analysis methods, such as spectral analysis (SA) (Zheng et al. 2010; Zhao et al. 2021), which decomposes time series data into different frequency components to analyze and predict their periodicity and frequency characteristics. However, the SA method necessitates a long-term and stable clock bias sequence, which is not feasible for the long-term forecasting of SCB due to its non-stationary long-term change (Heo et al. 2010). The other category comprises time domain analysis methods, including quadratic polynomial (QP) (Huang et al. 2011; Huang et al. 2018), GM(1,1) (Lu et al. 2008; Liang et al. 2016), autoregressive integrated moving average (ARIMA) (Xu et al. 2009; Xi et al. 2014; Zhang et al. 2022), Kalman Filtering model (Davis et al. 2012; Zhu et al. 2008), exponential smoothing method (Wang et al. 2017; Yu et al. 2020), and machine learning methods (Lu et al. 2023). QP, based on atomic clocks' physical properties, offers a clear physical interpretation but exhibits increasing prediction errors with longer forecast horizons and lacks the ability to capture clock bias' nonlinear dynamic characteristics (Jonsson and Eklundh 2002; Cui and Jiao 2005; Lei et al. 2022). The GM(1,1) model requires minimal sample data, simple modeling, and demonstrates good long-term forecasting performance, yet it necessitates a smooth and exponential change pattern in clock bias series, limiting its applicability (Lu et al. 2008). ARIMA can identify trends, seasonality, and cyclicity in time series data for future predictions but requires smooth data, making it unsuitable for long-term forecasting (Lei et al. 2022). The Kalman Filtering model predicts clock bias by recursively updating the state and calculating filtering results using system state and observation models. While effective in handling noise and uncertainty, it struggles with nonlinear problems and long-term dependencies, requiring accurate determination of process and observation noise to prevent errors in state estimation and reduce prediction accuracy (Lei et al. 2022). Exponential smoothing is typically suitable for short- and medium-term forecasts but may underperform with data exhibiting long trends or complex nonlinear relationships (Lei et al. 2022). Machine learning methods for clock bias prediction encompass support vector machine (SVM) (He et al. 2019), BP neural network (Bai et al. 2023), wavelet neural network (WNN) (Ai et al. 2016; Wang et al. 2017; Wang et al. 2020; Wang et al. 2021), radial basis function (RBF) (Wang et al. 2014), LSTM (Huang et al. 2021; He et al. 2023; Cai et al. 2024), and combination models like ARMA model and RBF neural network (Li et al. 2013), and wavelet transform and SVM (Lei et al. 2014). Neural networks, with rapid learning and superior generalization capabilities, are ideal for nonlinear time series forecasting, with ongoing research focusing on optimizing parameters to enhance forecasting accuracy and stability. However, challenges such as local optimization, overfitting, and underfitting persist, necessitating further exploration in data preprocessing, model parameter selection, and training strategies.

The SCB is characterized as a non-smooth and non-linear time series. Due to potential issues in the original SCB data, such as data loss, clock jumps, and gross errors (Feng 2009), utilizing the raw aberrant data quality may compromise the development and effectiveness of prediction models. Consequently, it is imperative to preprocess the initial SCB data post-acquisition to ensure the prediction model's reliability. Preprocessing techniques for the original SCB data typically involve methods such as the Median Absolute Deviation (MAD)

(Feng 2009), the 3σ criterion, and the enhanced MAD method (Wang et al. 2016; Huang et al. 2022), etc.. This study employs the MAD method for data preprocessing.

Previous research on predicting SCB has primarily focused on time domain analysis, neglecting the essential frequency domain information of the time series. To enhance prediction accuracy and stability by leveraging the frequency characteristics of SCB data, this study proposes utilizing Variational Mode Decomposition (VMD) (Dragomiretskiy and Zosso 2014) to decompose the SCB time series into multiple mode components. VMD is a signal decomposition technique suitable for data with non-linear and non-smooth attributes like SCB. By decomposing the data into distinct frequency components and trend information, VMD enables a better understanding of underlying patterns in the data, such as long-term trends, seasonal variations, and cyclical fluctuations. Notably, VMD can effectively isolate noise or disturbances as low-frequency mode components, facilitating clearer observation and analysis of true trends and cyclical changes in the data, thereby enhancing prediction accuracy.

Long Short-Term Memory (LSTM) represents a specialized form of recurrent neural network (RNN) distinguished by its superior memory capabilities and capacity to model extended dependencies (Yu et al. 2019). This is attributed to its distinctive architecture, which addresses the limitations of conventional RNNs in capturing lengthy sequences and effectively managing issues like gradient vanishing or explosion. Notably, LSTM excels in encoding sequential data, rendering it well-suited for addressing time series tasks such as SCB prediction. Despite the widespread application of LSTM in various domains for time series forecasting, the existing literature on SCB prediction utilizing LSTM remains relatively scarce (Huang et al. 2021; He et al. 2023; Cai et al. 2024), indicating a requirement for further investigation.

This study represents the inaugural endeavor to integrate VMD with LSTM for SCB prediction. Initially, the original SCB data undergoes phase-frequency conversion, followed by preprocessing the SCB frequency data using the MAD method. Subsequently, the preprocessed data is decomposed utilizing VMD, with the decomposed data components fed into the LSTM network. The predicted values of each component are then aggregated and reconstructed to derive the final SCB prediction. The efficacy and benefits of the VMD-LSTM model in the realm of SCB prediction are validated through experimental analysis. The results of the experiments demonstrate notable enhancements in SCB prediction accuracy when applied to the spaceborne hydrogen atomic clock for BeiDou Navigation Satellite System (BDS). Specifically, the model's clock bias prediction results for 6 hours, 3 days, and 15 days are 0.048 ns, 0.204 ns and 1.397 ns, respectively, showcasing superior performance compared to both the standalone LSTM network and BP neural network. The model's predictions exhibit a 56%, 84% and 83% improvement for 6 hours, 3 days, and 15 days, respectively, over the LSTM network, and a 59%, 82% and 83% enhancement over the BP neural network. By enhancing the precision and reliability of SCB prediction, this research aims to establish a robust foundation for applications such as precise navigation, positioning, timekeeping, time synchronization, and frequency control.

2 Variational Mode Decomposition

The technique known as VMD was introduced by K. Dragomiretskiy et al. (2014) as a signal processing method. VMD is capable of breaking down a multi-component signal into

individual amplitude modulation (AM) and frequency modulation (FM) signals simultaneously. Through an iterative process of optimizing a variational model, VMD identifies the frequency center and bandwidth of each signal component. This enables the adaptive dissection of the signal in the frequency domain and the effective isolation of each component. Specifically, if the original signal comprises *K* sub-signals with distinct center frequencies and finite bandwidths (referred to as Intrinsic Mode Functions, IMFs), VMD utilizes the Alternating Direction Multiplier Method (ADMM) and the Parseval theorem to iteratively update the center frequency and bandwidth of each mode. Subsequently, the modes are demodulated to their respective fundamental bands, resulting in the extraction of mode components with varying center frequencies.

The essence of the VMD technique lies in formulating a variational problem and subsequently solving it.

(a) Intrinsic Mode Function

The VMD method redefines IMF within a limited frequency range by imposing stricter limitations, as formally represented through mathematical expressions:

$$u_k(t) = A_k(t)\cos(\phi_k(t)) \tag{2-1}$$

In the given equation, $A_k(t)$ represents the envelope amplitude of signal $u_k(t)$, while $\phi_k(t)$ denotes the instantaneous phase.

(b) Development of Variational Problem

The variational problem is subject to two primary constraints: firstly, the total bandwidth of the central frequencies of each modal component must be minimized; and secondly, the sum of all modal components must be equivalent to the original signal.

$$\begin{cases} \min\limits_{\{u_k\},\{\omega_k\}} \left\{ \sum\limits_k \left\| \partial_t \left[\left(\delta(t) + \dfrac{j}{\pi t} \right) * u_k(t) \right] e^{-j\omega_k t} \right\|_2^2 \right\} \\ \text{s.t.} \sum\limits_k u_k = f \end{cases} \tag{2-2}$$

where $u_k = \{u_1, u_2, \ldots, u_K\}$ represents the collection of mode functions, $\omega_k = \{\omega_1, \omega_2, \ldots, \omega_K\}$ denotes the ensemble of center frequencies, ∂_t signifies the partial derivative of the function in relation to the variable time *t*, $\delta(t)$ stands for the Dirac function, and the symbol * denotes the convolution operator.

To minimize the bandwidth between individual modes derived from VMD and ensure their sum equals the original signal *f*, a series of steps are undertaken. Firstly, the Hilbert transform is applied to each mode function $u_k(t)$ to obtain the unilateral spectrum $\left(\delta(t) + \dfrac{j}{\pi t} \right) * u_k(t)$. Subsequently, this spectrum is adjusted by multiplying it with the estimated center frequency $e^{-j\omega_k t}$ to align each mode's spectrum with the estimated center frequency. Finally, the bandwidth of each mode's signal is determined by employing Gaussian smoothing, specifically by calculating the square root of the gradient of the L_2 paradigm.

(c) Addressing of sVariational Problem

To determine the optimal solution for the variational constraint model described above, the Lagrange multiplier λ and the quadratic penalty factor α are incorporated to convert the inequality constraints into equational constraints. This transformation yields an augmented Lagrange expression:

$$L(\{u_k\},\{\omega_k\},\lambda)=\alpha\sum_{k=1}^{K}\left\|\partial_t\left[\left(\delta(t)+\frac{j}{\pi t}\right)*u_k(t)\right]e^{-j\omega_k t}\right\|_2^2+\left\|f(t)-\sum_{k=1}^{K}u_k(t)\right\|_2^2+\left\langle\lambda(t),f(t)-\sum_{k=1}^{K}u_k(t)\right\rangle \tag{2-3}$$

where α serves to mitigate the effects of Gaussian noise, while λ ensures the robustness of the constraint issue.

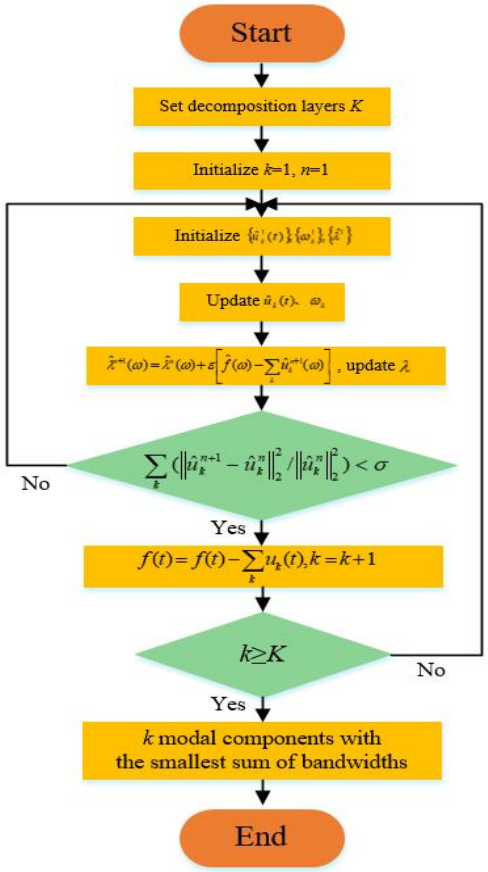


Fig.1 VMD Flowchart

As illustrated in Fig. 1, the VMD technique is executed through the subsequent procedures.

(i) Initialize $\{\hat{u}_k^1\}$, $\{\omega_k^1\}$, $\{\hat{\lambda}^1\}$, $n\leftarrow 0$;

(ii) Set the number of iterations $n=n+1$;

(iii) When $\omega\geq 0$, update the spectrum $\hat{u}_k$ of each mode accordingly:

$$\hat{u}_k^{n+1}(\omega)=\frac{\hat{f}(\omega)-\sum_{i=1,i<k}\hat{u}_i^{n+1}(\omega)-\sum_{i=1,i>k}\hat{u}_i^{n}(\omega)+\frac{\hat{\lambda}^n(\omega)}{2}}{1+2\alpha(\omega-\omega_k^n)^2}$$

Update the center frequency ω_k:

$$\omega_k^{n+1}(\omega)=\frac{\int_0^\infty\omega\left|\hat{u}_k^{n+1}(\omega)\right|^2\mathrm{d}\omega}{\int_0^\infty\left|\hat{u}_k^{n+1}(\omega)\right|^2\mathrm{d}\omega}$$

(iv) Update the Lagrange multiplier λ:

$$\hat{\lambda}^{n+1}(\omega)=\hat{\lambda}^n(\omega)+\varepsilon\left[\hat{f}(\omega)-\sum_k\hat{u}_k^{n+1}(\omega)\right]$$

Among which, ε is the noise tolerance parameter. For signals with strong noise, setting ε to 0 can achieve better noise reduction effects.

(v) Continue to perform Steps (ii) through (iv) repeatedly until the specified iterative constraints are met:

$$\sum_k (\left\|\hat{u}_k^{n+1} - \hat{u}_k^n\right\|_2^2 / \left\|\hat{u}_k^n\right\|_2^2) < \sigma$$

The variable σ represents the precision of the discriminant, and $\sigma > 0$.

3 Long Short-Term Memory Network

The Long Short-Term Memory (LSTM) network is an enhanced recurrent neural network (RNN) algorithm introduced by S. Hochreiter and J. Schmidhuber (1997). It incorporates specialized gating mechanisms to regulate the information flow and memory updating processes. These gating units, such as the input gate, forget gate, and output gate, utilize a combination of sigmoid activation functions and dot-multiplication operations to selectively retain or discard information. This selective retention enables the LSTM network to focus on pertinent information for making predictions, mirroring the cognitive processes of the brain in extracting and articulating key concepts.

As illustrated in Fig.2, the symbols *i*, *f*, and *o* are utilized to represent the input gate, forget gate, and output gate, correspondingly. The symbol $\otimes$ signifies the element-wise multiplication, while $\oplus$ denotes the element-wise addition. The weight matrix and bias vector of the network are represented by *W* and *b*, respectively. At time *t*, the input and output vectors of the LSTM hidden layer are denoted as x_t and h_t, respectively. The memory cell is denoted as C_t, with C_{t-1} representing the preceding memory cell state. The outputs of the forget gate and input gate are represented by f_t and i_t, respectively, while the candidate information is denoted as $\tilde{C}_t$. Specifically, σ represents the sigmoid function, and tanh signifies the hyperbolic tangent function.

$$\sigma(x) = \frac{1}{1+e^{-x}}$$

$$\tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}}$$

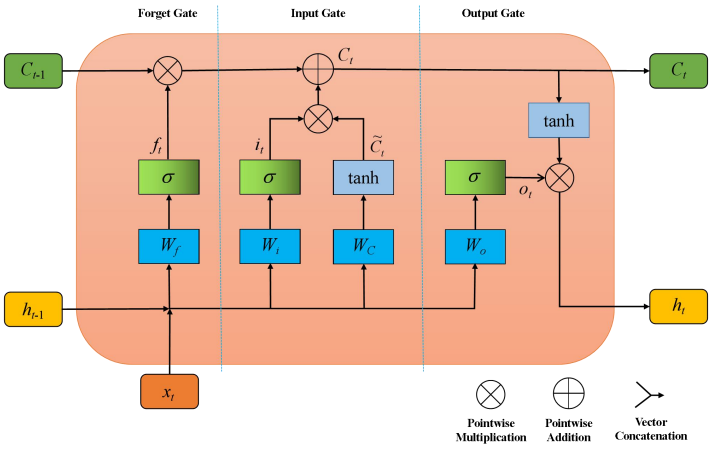


Fig.2 LSTM cell and Network Architectures

(a) Input Gate

The input gate is comprised of a sigmoid activation function and a hyperbolic tangent tanh activation function. The sigmoid function produces an output ranging from 0 to 1, determining the significance of information. On the other hand, the tanh activation function generates new candidate information denoted as $\widetilde{C}_t$ within the range of -1 to 1. The product of the output i_t from the input gate and the candidate information $\widetilde{C}_t$ is calculated, and this result is utilized in the updating process of the memory cell.

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i)$$
$$\widetilde{C}_t = \tanh(W_C[h_{t-1}, x_t] + b_C)$$

(b) Forget Gate

The forget gate is comprised solely of a sigmoid activation function that generates a value ranging from 0 to 1, determining the retention or discarding of memories. Additionally, it effectively mitigates the issues of vanishing and exploding gradients encountered during backward gradient propagation over time. The sigmoid function's output is directly applied to the previous memory cell state, preserving information with outputs closer to 1 and discarding information with outputs closer to 0.

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f)$$

(c) Output Gate

The output gate plays a crucial role in determining the specific information from the memory cell that is to be transmitted to the hidden state at the present moment, thereby selecting the segment of the memory cell to be conveyed at time *t*. This gate comprises both a sigmoid activation function and a tanh activation function. The sigmoid function is responsible for determining the pertinent information for output, while the tanh function is involved in processing the memory cell's state in readiness for output. The output from the sigmoid function is then multiplied by the memory cell's state that has been processed by the tanh function, resulting in the generation of the hidden state at the current time.

$$C_t = f_t \otimes C_{t-1} + i_i \otimes \widetilde{C}_t$$
$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$$
$$h_t = o_t \otimes \tanh(C_t)$$

The LSTM network is comprised of one or more LSTM cells, with a deep network being formed by stacking multiple LSTM cells. Initially, the output of each neuron is computed in a forward manner, followed by the calculation of the error term for each neuron through back-propagation of errors. Subsequently, the weight matrix and parameters are adjusted based on the principles of gradient descent. By iteratively updating the weight matrix and parameters within the LSTM using sample data, the network's overall error is minimized until it meets a predefined threshold or the maximum number of training iterations is reached, signifying the completion of LSTM training. The trained LSTM model can then be utilized for data prediction.

4 SCB Prediction Model Construction based on VMD-LSTM

As shown in Fig.3, the VMD-LSTM prediction model outlined in this study primarily comprises the subsequent procedures:

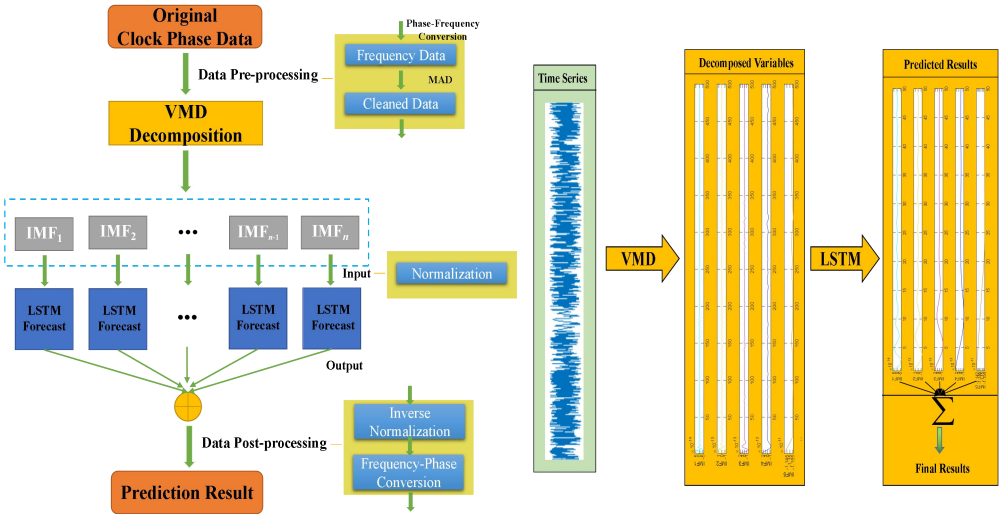


Fig.3 SCB Prediction Model based on VMD-LSTM

(a) Data Preprocessing

Initially, the time series data acquired undergo preprocessing to address potential issues within the original SCB dataset, such as missing data, clock jumps, and gross errors. Utilizing SCB data with quality concerns directly may compromise the development and effectiveness of the SCB prediction model. Therefore, it is imperative to preprocess the raw SCB data to ensure the reliability of the prediction model. Furthermore, the original SCB data is in phase form, typically featuring large values (typically ranging from E-4 to E-5), making it challenging to identify outliers directly. Consequently, the SCB phase data is transformed into frequency data, and any outliers within the SCB frequency data are addressed. This study employs the MAD method for outlier processing within the data.

(b) Data Partitioning

Following the removal of outliers, the data is partitioned into three distinct sets: the training set, validation set, and test set, as described in Fig.4. The training set is employed for model training purposes, the validation set aids in fine-tuning hyperparameters and model selection, while the test set is utilized to assess the overall performance and generalization capabilities of the final model.

Fig.4 Training Set, Validation Set and Test Set

(c) Series Decomposition

The time series data is analyzed through the VMD technique to extract multiple intrinsic mode components IMF_i and a residual term. VMD requires the specification of parameters such as the mode number *K*, penalty factor α, and convergence stopping condition σ. The challenge and critical aspect of the process involve the selection of appropriate values for *K* and α, while σ is typically set to a default value. Opting for a higher mode number may result in mode repetition or additional noise, whereas a lower number could lead to underdecomposition of the modes. The parameter α primarily influences the bandwidth of the IMF_i, with larger values narrowing the bandwidth of each IMF_i component, and conversely, smaller values widening the bandwidth of the component.

In this study, the *K* value is ascertained through the observation center frequency technique. The predetermined mode number *K* is systematically varied from small to large, and the optimal value of *K* is considered to be reached when the center frequency of the final layer of the IMF_i component exhibits a consistent stability.

(d) Data Normalization

Data normalization involves the process of standardizing the data by mapping the input and output features of a sample dataset to a range of [-1,1]. This procedure aims to mitigate the impact of varying magnitudes within the data, thereby enhancing the convergence rate of the model.

(e) Prediction Using LSTM

Distinct LSTM network models are created for individual mode components IMF_i to forecast outcomes. The optimization of hyperparameters plays a crucial role in enhancing the predictive precision of the LSTM network. The hyperparameters considered in the model outlined in this study encompass the quantity of layers, the dimensions of the hidden layer, the batch size, the random dropout rate, the epoch, the loss function, the optimizer, the learning rate, the training step, the prediction step, and the number of training iterations, etc.. To mitigate the risk of error accumulation, single-step prediction methodology is adopted in this investigation.

(f) Prediction Component Reconstruction

The anticipated values of individual mode components are combined and aggregated, followed by inverse normalization and frequency-phase conversion processing to derive the ultimate prediction outcomes for the SCB.

(g) Model Evaluation Metrics

This study employs five metrics to assess the efficacy of the SCB prediction model

utilizing VMD-LSTM. These metrics include Mean Absolute Error (MAE), Mean Square Error (MSE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), and R-Square (R-2).

Suppose the predicted values $\hat{y} = \{\hat{y}_1, \hat{y}_2, \ldots, \hat{y}_n\}$ and the true values $y = \{y_1, y_2, \ldots, y_n\}$.

MAE signifies the average absolute error and provides insight into the accuracy of the predicted values, and the greater the error, the greater the value.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\hat{y}_i - y_i\right|$$

MSE represents the ratio of the sum of squares of the deviation from the predicted and actual values to the total number of samples.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2$$

RMSE is utilized to determine the square root of the sum of squared deviations between predicted and actual values relative to the total sample size, and the greater the error, the greater the value.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}$$

MAPE can eliminate the impact of dimension and be used to evaluate bias objectively.

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right|$$

Additionally, R2, a measure of goodness of fit, is employed to evaluate the degree to which the fitted values align with the observed data. Generally speaking, the larger the R-Square, the better the model fitting effect.

$$R2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(\bar{y}_i - y_i)^2}$$

5 Case Analysis

This study utilizes the 300-second precision SCB product of BDS provided by the International GNSS Monitoring and Assessment Service (IGMAS) for a duration of 106 days spanning from Sep. 6th to Dec. 20th, 2023 (denoted as 20230906-20231220). In order to evaluate the accuracy and universality of SCB prediction using VMD-LSTM model, this paper divides these 106-day data into five segments every 5 days, namely, 20230906-20231130 (Segment 1), 20230911-20231205 (Segment 2), 20230916-20231210 (Segment 3), 20230921-20231215 (Segment 4), and 20230926-20231220 (Segment 5). The

training of the model involves using data from the initial 70 days as the training set and data from the subsequent day as the validation set. SCB predictions are conducted for varying durations of 6 hours, 3 days and 15 days.

The spaceborne atomic clocks utilized by BDS in orbit predominantly consist of Rubidium (Rb) and Hydrogen (H) atomic clocks, with specific types of atomic clocks currently operational in orbit detailed in Table 1.

Table 1 Spaceborne Atomic Clock Types of BDS

Types	**PRN**
GEO	01(Rb) 02(Rb) 03(Rb) 04(Rb) 05(Rb) 59(H) 60(H)
IGSO	06(Rb) 07(Rb) 08(Rb) 09(Rb) 10(Rb) 16(Rb) 38(H) 39(H) **40(H)**
	11(Rb) 12(Rb) 13(Rb) 14(Rb) 19(Rb) 20(Rb) 21(Rb) 22(Rb) 23(Rb)
MEO	24(Rb) 25(Rb) 26(H) **27(H)** 28(H) 29(H) 30(H) 32(Rb) 33(Rb) 34(H)
	35(H) 36(Rb) 37(Rb) 41(Rb) 42(Rb) 43(H) 44(H) 45(H) 46(H)

In this study, the VMD-LSTM model is utilized to investigate the prediction performance of clock bias in hydrogen clocks within BDS. As a result of the limited availability of GEO data pertaining to BDS provided by IGMAS during the specified timeframe, two satellites PRN27 (H) and PRN40 (H) in inclined geosynchronous orbit (IGSO) and medium Earth orbit (MEO) were randomly selected for experimental analysis. An illustration of PRN27 (Segment 1) is provided below for reference.

5.1 Data Preprocessing of PRN27

The initial phase data of PRN27 during the timeframe from Segment 1 is presented as Fig.5 below.

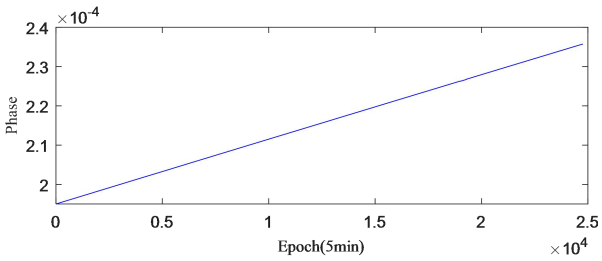


Fig.5 Atomic Clock Phase Data of PRN27

As depicted in Fig.5, it is evident that the outliers are not readily discernible from the initial phase data. Consequently, a phase-frequency conversion process is conducted to derive the atomic clock frequency data, as illustrated in Fig.6.

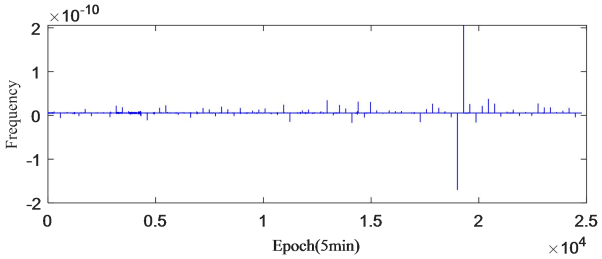


Fig.6 Atomic Clock Frequency Data of PRN27

In the atomic clock frequency data presented in Fig.6, certain data points that deviate significantly from the overall trend can be readily distinguished. In this study, these outliers

are eliminated through the employment of MAD (set the threshold factor equal to 3), whereby the outliers are replaced through linear interpolation. Subsequently, Fig.7 displays the frequency data pertaining to PRN27 subsequent to the removal of outliers.

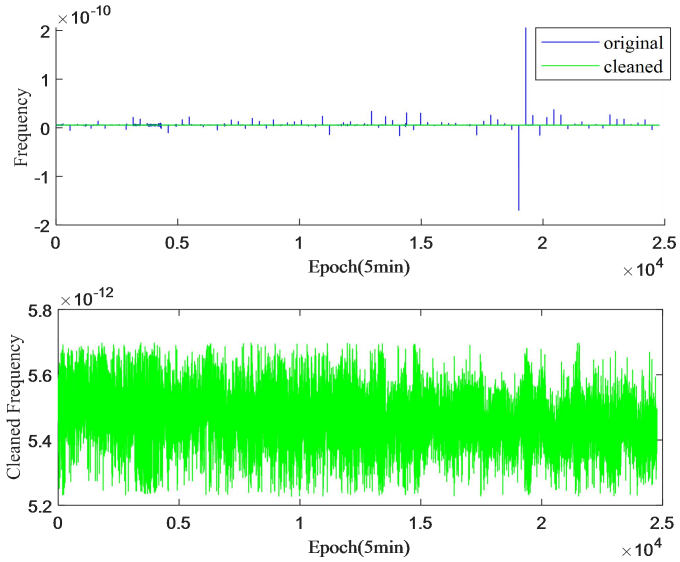


Fig.7 Atomic Clock Frequency Data of PRN27 after Preprocessing

5.2 Decomposition Using VMD

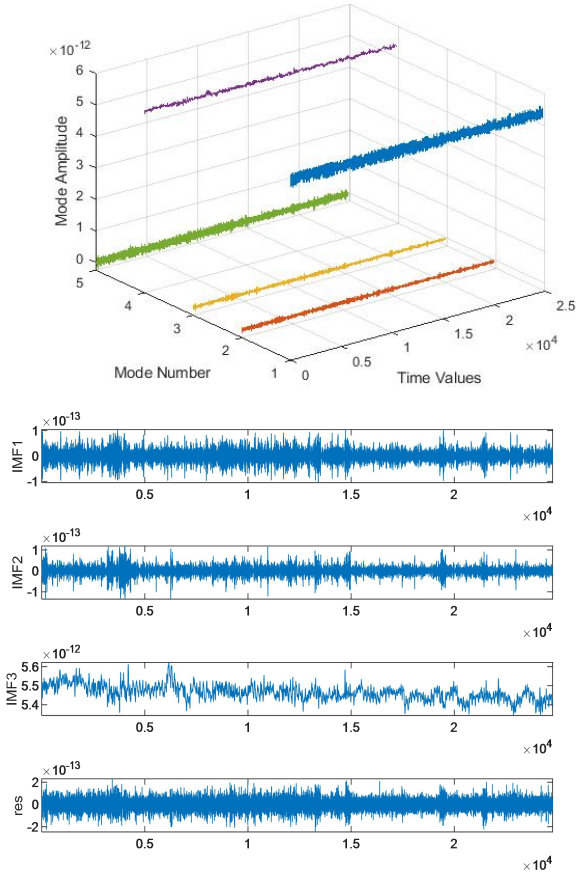


Fig.8 Frequency Data Decomposition for PRN27 Using VMD

Here, VMD is utilized to ascertain the quantity of sequence decompositions, with a total of three sequence decompositions selected for analysis. The decomposition outcomes of VMD applied to the atomic clock frequency data pertaining to PRN27 are illustrated in Fig.8.

The figure displays the original waveform in the first column, while columns 2-4 represent the decompositions from high-frequency to low-frequency sequences. The final column exhibits the residual component of the decomposition process.

5.3 Prediction Using LSTM

Following the decomposition of the atomic clock frequency data through VMD, the data is subsequently normalized for each component and fed into the LSTM network model. The hyperparameters for the LSTM model utilized in this study were determined as outlined below:

Table 2 Hyperparameter Selection for LSTM Network Model

No	Parameter	Value
1	Loss Function	MSE
2	Optimizer	Adam
3	Number of LSTM Layers	[32,32]
4	Number of Training Rounds	30150
5	Input Dimension Size	10
6	Output Dimension Size	1
7	Batchsize	100
8	Learning Rate	0.0002
9	Epoch	150

The frequency prediction performance of PRN27 for 15 days are presented in Fig.9 as follows:

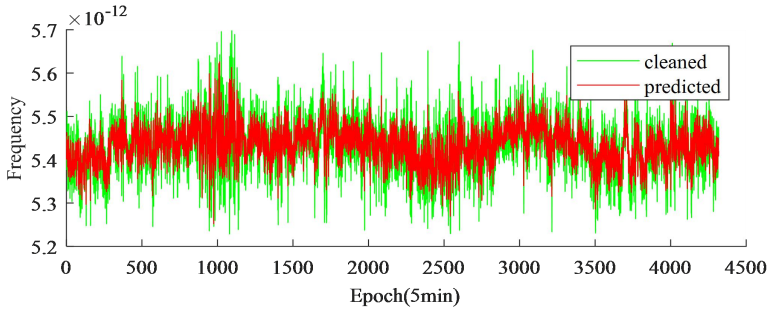


Fig.9 Frequency Prediction Performance of PRN27 for 15 days

5.4 SCB Prediction Results

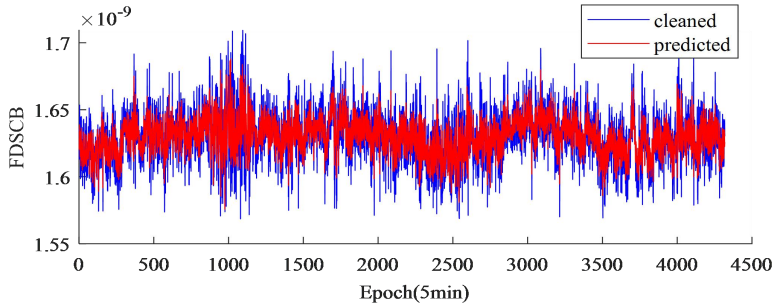


Fig.10 The cleaned and predicted FDSCB of PRN27 for 15 days

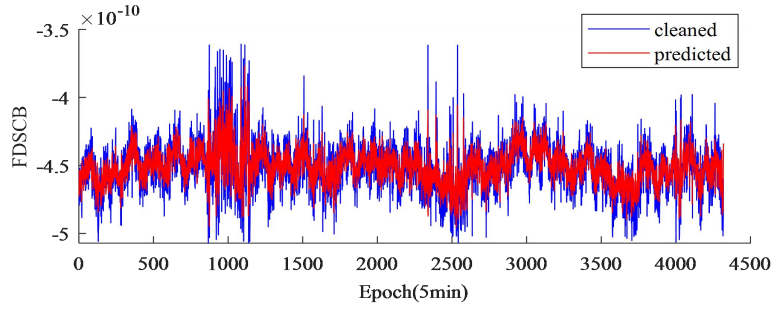


Fig.11 The cleaned and predicted FDSCB of PRN40 for 15 days

Upon the completion of the frequency prediction, the predicted frequency information is transformed into phase data (predicted), with the prediction outcome of SCB being determined by the RMSE between the cleaned and predicted phase data. The first difference of SCB (FDSCB) for PRN27 and PRN40 for 15 days in Segment 1 are shown in Fig.10 and Fig.11, where the the blue curve is the true value (cleaned) of FDSCB, and the red one is the corresponding predicted value.

Table 4 illustrates the clock bias prediction outcomes for two BDS satellites, namely PRN27 and PRN40, utilizing VMD-LSTM methodology for Segments 1 to 5.

Table 4. SCB Prediction Results (RMSE) of PRN27 and PRN40 Using VMD-LSTM (ns)

Time	PRN27 (MEO)			PRN40 (IGSO)		
	6 hours	3 days	15 days	6 hours	3 days	15 days
Segment 1	0.082	0.426	1.202	0.067	0.599	1.795
Segment 2	0.026	0.113	0.629	0.047	0.194	0.477
Segment 3	0.029	0.163	0.828	0.036	0.152	0.938
Segment 4	0.048	0.069	0.645	0.100	0.231	1.458
Segment 5	0.029	0.037	1.902	0.011	0.048	4.090
Average (Segments1-5)	**0.043**	**0.162**	**1.041**	**0.052**	**0.245**	**1.752**

5.5 Comparison and Analysis

To evaluate the efficacy of SCB prediction utilizing the VMD-LSTM model, this study employs the LSTM model and the BP neural network model to forecast the identical dataset. The outcomes are presented in Table 5 and Table 6 as follows:

Table 5. SCB Prediction Results (RMSE) of PRN27 and PRN40 Using LSTM (ns)

Time	PRN27 (MEO)			PRN40 (IGSO)		
	6 hours	3 days	15 days	6 hours	3 days	15 days
Segment 1	0.350	2.553	8.875	0.192	1.845	9.620
Segment 2	0.108	0.771	7.557	0.050	1.287	7.468
Segment 3	0.057	0.552	6.432	0.087	0.695	6.432
Segment 4	0.062	0.828	7.799	0.065	1.500	8.429
Segment 5	0.056	1.583	9.437	0.057	0.991	10.840
Average (Segments1-5)	**0.127**	**1.257**	**8.020**	**0.090**	**1.264**	**8.558**

Table 6. SCB Prediction Results (RMSE) of PRN27 and PRN40 Using BP (ns)

Time	PRN27 (MEO)			PRN40 (IGSO)		
	6 hours	3 days	15 days	6 hours	3 days	15 days
Segment 1	0.255	2.366	9.557	0.221	1.772	8.276
Segment 2	0.025	0.650	8.033	0.151	1.221	8.027
Segment 3	0.043	0.494	6.462	0.091	0.607	6.059
Segment 4	0.068	0.588	8.316	0.188	1.209	5.523
Segment 5	0.031	1.281	10.344	0.091	0.997	12.715
Average (Segments1-5)	**0.084**	**1.076**	**8.542**	**0.148**	**1.161**	**8.120**

We use the average value of five different segments to represent the prediction results of SCB using the three methods, and the RMSE comparison is presented in Fig.12.

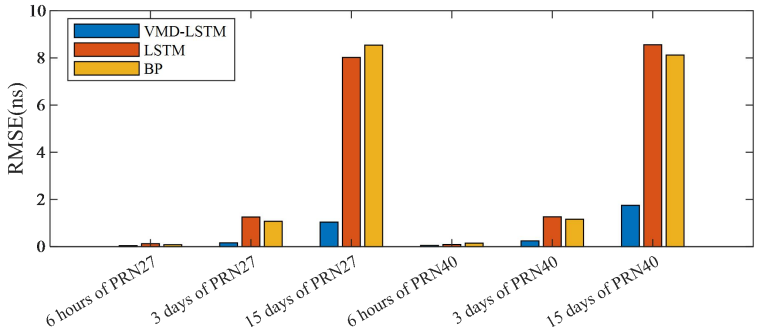
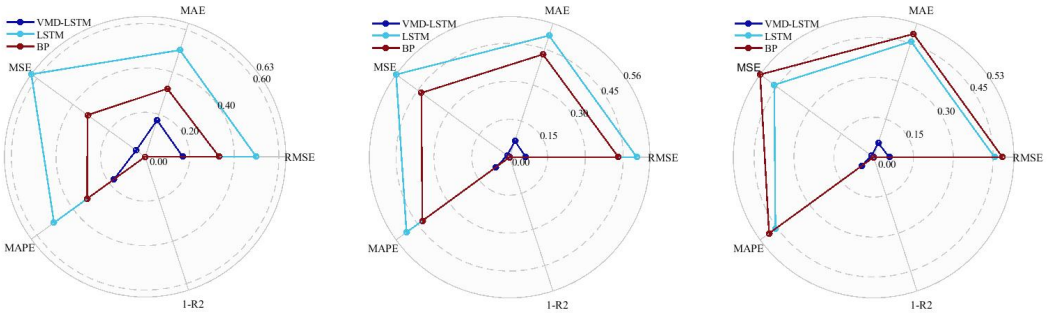


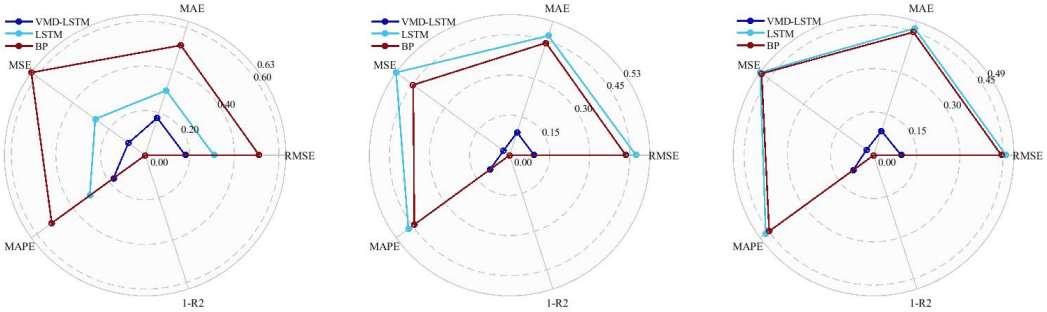
Fig.12 RMSE Comparison of PRN27 and PRN 40 for 6 hours, 3 days and 15 days

The detail comparison results of the evaluation metrics of the VMD-LSTM, LSTM, and BP neural network prediction models are shown in Fig.13 and Fig.14. To facilitate comparison, we normalize each metric into a single dimension. Obviously, each metric of the SCB prediction model based on VMD-LSTM is optimal.



(a) 6 hours for PRN27 (b) 3 days for PRN27 (c) 15 days for PRN27

Fig.13 Metrics Comparison of PRN27 for 6 hours, 3 days and 15 days



(a) 6 hours for PRN40 (b) 3 days for PRN40 (c) 15 days for PRN40

Fig.14 Metrics Comparison of PRN40 for 6 hours, 3 days and 15 days

This paper takes the average prediction results of PRN27 and PRN40 satellites for 6 hours, 3 days, and 15 days as the predicted results of each model for the SCB of BDS hydrogen atomic clock. The analysis reveals that, under the VMD-LSTM model, the predicted SCB values for 6 hours, 3 days and 15 days are 0.048 ns, 0.204 ns and 1.397 ns, respectively. In contrast, the LSTM model predicts SCB values of 0.109 ns, 1.261 ns and 8.289 ns for the same time intervals. Similarly, the BP neural network model forecasts SCB values of 0.116 ns, 1.119 ns and 8.331 ns, respectively.

Comparing the results presented in Tables 4-6 and Figs.11-13, it is evident that the SCB predictions derived from the VMD-LSTM model exhibit superior performance compared to those from the LSTM and BP neural network models. Specifically, the VMD-LSTM model's

predictions for 6 hours, 3 days and 15 days are approximately 56%, 84% and 83% higher than the LSTM model, and 59%, 82% and 83% higher than the BP neural network model.

There are two reasons why the VMD-LSTM-based SCB prediction model established in this paper is effective: Firstly, transforming phase data into frequency data for prediction makes the prediction results closer to the true values. We verified this point by directly predicting phase data after removing outliers using VMD-LSTM, but the effect was not as good as predicting with frequency data and then converting them into phase data. The possible reason is that the phase-frequency conversion is equivalent to a low-pass filter, which has the effect of filtering out noise. Secondly, signal decomposition improves the prediction accuracy. By decomposing the signal into multiple mode components using VMD and taking them as inputs for the LSTM model, the local features and temporal-spatial relationships of the signal can be better captured, thereby improving the model's prediction ability and generalization ability.

6. Conclusions

The precise estimation of the clock bias in atomic clocks holds significant importance for accurate timekeeping, navigation and positioning. This study introduces a novel approach for forecasting atomic clock bias by integrating signal decomposition and deep learning techniques, specifically utilizing VMD in conjunction with LSTM neural networks. Empirical findings demonstrate that for the BDS satellite hydrogen atomic clock, the projected clock errors over 6 hours, 3 days and 15 days are 0.048 ns, 0.204 ns and 1.397 ns, respectively, showcasing notable enhancements compared to both the LSTM and BP neural network models.

Decomposing time series data can enhance prediction accuracy by enabling a more comprehensive understanding of the data components, facilitating the selection of suitable forecasting methods based on trends, seasonality, and residuals. Trend analysis aids in forecasting long-term patterns, seasonality analysis predicts periodic fluctuations, and residual analysis identifies anomalies or enhances forecasting models. VMD is a signal processing technique used to decompose complex time series signals into multiple local frequency patterns that can reflect different frequency components and amplitude variations in the signal. The main idea of VMD is to decompose the signal into multiple IMFs, which have different frequency and amplitude characteristics and are orthogonal to a certain extent, through optimization problems. Compared with traditional signal decomposition methods, VMD can better adapt to nonlinear and non-stationary signals, and has better local feature extraction capability. Consequently, employing an LSTM model to predict each data component leads to outcomes that closely align with actual values. By enhancing the precision and reliability of atomic clock error predictions, this study aims to establish a robust foundation for precise navigation, positioning, timekeeping, time synchronization, frequency regulation, and related applications. Future research will focus on optimizing LSTM model parameters and extending its application within the realm of time establishment and timekeeping.

Acknowledgements

Thanks Prof. Huang Guanwen of Chang'an University, Prof. Liang Kun of Beijing Jiaotong University, Prof. Wang Bo of Tsinghua University and Dr. Wang Yupu of Beijing Satellite

Navigation Center for their discussion and guidance.
Thanks to International GNSS Monitoring & Assessment System (http://www.igmas.org) for providing data support.

Author Contributions

SWL proposed the method used in this paper and complete the experiment; SFF, LYT and XJF discussed and analyzed the experimental results and the manuscript, and proposed suggestions for improvements; HSF, WGC and MP discussed and analyzed the data preprocessing of the experimental process and the experimental results; LYX and ZR reviewed the entire manuscript and performed proofreading. All authors read and approved the final manuscript.

Funding

There is no any funding or project support for this manuscript.

Data availability

The datasets analyzed during the current study are available in the International GNSS Monitoring & Assessment System (http://www.igmas.org). You can also contact the corresponding author of this manuscript to obtain all the experimental data and results.

Declarations

Competing Interests

The authors declare that they have no competing interests.

References

Ai QS, Xu TH, Li JJ, Xiong HW (2016) The short-term forecast of BeiDou satellite clock bias based on wavelet neural network. In: 2016 China Satellite Navigation Conference, pp.145-154. https://doi.org/10.1007/978-981-1090934-1_14

Bai HW, Cao QQ, An SB (2023) Mind evolutionary algorithm optimization in the prediction of satellite clock bias using the back propagation neural network. Scientific Reports 13:2095. https://doi.org/10.1038/s41598-023-28855-y

Cai CL, Liu MY, Li PC, Li ZX, Lv KH (2024) Enhancing satellite clock bias prediction in BDS with LSTM-attention model. GPS Solutions 28:92. https://doi.org/10.1007/s10291-024-01640-8

Cui XQ, J WH (2005) Grey system model for the satellite clock error predicting. Geomatics and Information Science of Wuhan University 30(5):447-450.

Davis J, Bhattarai S, Ziebart M (2012) Development of a Kalman filter based GPS satellite clock time-offset prediction algorithm. In: 2012 IEEE European Frequency and Time Forum, pp,152-156. https://doi.org/10.1109/EFTF.2012.6502355

Dragomiretskiy K, Zosso D (2014) Variational mode decompostion. In: IEEE Transactions on Signal Processing 62(3):531-544. https://doi.org/10.1109/TSP.2013.2288675

Feng SL (2009) Study on the methods of data preprocessing and performance analysis for atomic clocks. Zhengzhou: Information Engineering University.

He LN, Zhou HR, Liu ZQ, Wen YL, H XF (2019) Improving clock prediction algorithm for BDS-2/3 satellites based on LS-SVM method. Remote Sensing 11(21):2554. https://doi.org/10.3390/rs11212554

He SF, Liu JL, Zhu XW, Dai ZQ, Li D (2023) Research on modeling and predictions of BDS-3 satellite clock bias using the LSTM neural network model. GPS Solutions 27(3):108.

https://doi.org/10.1007/s10291-023-01451-3
Heo YJ, Cho J, Heo MB (2010) Improving prediction accuracy of GPS satellite clocks with periodic variation behaviour. Measurement Science and Technology 21(7):073001. Https://doi.org/10.1088/0957-0233/21/7/073001
Hochreiter S, Schmidhuber J (1997) Long short-term memory. Neural Computation 9(8):1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735
Huang BH, Ji ZX, Zhai RJ, Xiao CF, Yang F, Yang BH, Wang YP (2021) Clock bias prediction algorithm for navigation satellites based on a supervised learning long short-term memory neural network. GPS Solutions 25(2):80. https//doi.org/10.1007/s10291-021-01115-0
Huang BH, Yang BH, Li MG, Guo ZK, Mao JY, Wang H (2022) An improved method for MAD gross error detection of clock error. Geomatics and information science of Wuhan University 47(5):747-752. https://doi.org/10.13203/j.whugis20190430
Huang GW, Cui BB, Zhang Q, Fu WJ, Li PL (2018) An improved predicted model for BDS ultra-rapid satellite clock offsets. Remote Sensing 10(1):60. https://doi.org/10.3390/rs10010060
Huang GW, Yang YX, Zhang Q (2011) Estimate and predict satellite clock error using adaptively robust sequential adjustment with classified adaptive factors based on opening windows. Acta Geodaetica Et Cartographica Sinica 40(1):15-21
IGS (2024). Products. http://www.igs.org/products, accessed 14th, May, 2024.
Jonsson P, Eklundh L (2002) Seasonality extraction by function fitting to time-series of satellite sensor data. IEEE Transactions on Geoscience Remote Sensing 40(8):1824-1832. https://doi.org/10.1109/TGRS.2002.802519
Lei Y, Xu JS, Cai HB (2022) Long-term prediction of satellite clock offset based on artificial neural network. Geomatics Science and Technology 10(3):131-139. https://doi.org/10.12677/gst.2022.103013
Lei Y, Zhao DN, Li B, Gao YP (2014) Prediction of satellite clock bias based on wavelet transform and least square support vector machines. Geomatics and Information Science of Wuhan University 39(7):815-819.
Li XY, Dong XR, Zheng K, Liu YT (2013) Research of satellite clock error prediction based on RBF neural network and ARMA model. In: 2013 China Satellite Navigation Conference, pp. 325-334. https://doi.org/10.1007/978-3-642-37407-4_29
Liang YJ, Ren C, Yang XF, Pang GF, Lan L (2016) A Grey model based on first difference in the application of the satellite clock bias prediction. Acta Astronom Sinica 56(3):264-277. https://doi.org/10.1016/j.chinastron.2016.01.008
Lu XF, Yang ZQ, Jia XL, Cui XQ (2008) Parameter optimization method of gray system theory for the satellite clock error predicting. Geomatics and Information Science of Wuhan University 33(5):492-495.
Lu YW, Zheng LQ, Hu C (2023) Analysis and comparison of satellite clock error prediction based on various deep learning algorithms. GNSS World of China 48(5):46-55, 91. https://doi.org/10.12265/j.gnss.2023138.
Wang GC, Liu LT, Xu AG, Su XQ, Liang XH (2014) The application of radial basis function neural network in the GPS satellite clock bias prediction. Acta Geodaetica Et Cartographica Sinica 43(8):803-807.
Wang L, Zhang Q, Huang GW, Tian J (2017) GPS satellite colck bias prediction based on

exponential smoothing method. Geomatics and Information Science of Wuhan University 42(7):995-1001. https://doi.org/10.13203/j.whugis20150089
Wang X, Chai HZ, Wang C (2020) A high-precision short-term prediction method with stable performance for satellite clock bias. GPS Solutions 24:105. https:doi.org/10.1007/s10291-020-01019-5
Wang X, Chai HZ, Wang C, Xiao GR, Chong Y, Guan XG (2021) Improved wavelet neural network based on change rate to predict satellite clock bias. Survey Review 53(379):325-334. https:doi.org/10.1080/00396265.2020.1758999
Wang YP, Lv ZP, Chen ZS, Huang LY (2014) A method of satellite clock bias prediction based on grey model and wavelet neural network. Journal of Geodesy and Geodynamics 34(3):155-159.
Wang YP, Lv ZP, Chen ZS, Huang LY, Li LY, Gong XC (2016) A new data preprocessing method of satellite clock bias and its application in WNN to predict medium-term and long-term clock bias. Geomatics and Information Science of Wuhan University 41(3):373-379. https://doi.org/10.13203/j.whugis20140216

Wang YP, Lu ZP, Qu YY, Li LY, Wang N (2017) Improving prediction performance of GPS satellite clock bias based on wavelet neural network. GPS Solutions 21(2):523-534.

Xi C, Cai CL, Li SM, Li XH, Li ZB, Deng KQ (2014) Long-term clock bias prediction based on an ARMA model. Chinese Astronomy and Astrophysics 38(3):342-354. https://doi.org/10.1016/j.chinastron.2014.07.010

Xu JY, Zeng AM (2009) Application of ARIMA (0,2,q) model to prediction of satellite clock error. Journal of Geodesy and Geodynamics 29(5):116-120.
Yu Y, Huang M, Wang CY, Hu R, Duan T (2020) A new BDS-2 satellite clock bias prediction algorithm with an improved exponential smoothing method. Applied Sciences 10(21):7456. https://doi.org/10.3390/app10217456
Yu Y, Si XS, Hu CH, Zhang JX (2019) A review of recurrent neural networks: LSTM cells and network architectures. Neural Computation 31(7):1235-1270. https://doi.org/10.1162/neco_a_01199
Zhang GC, Han SH, Ye J, Hao RZ, Zhang JC, Li X, Jia K (2022) A method for precisely predicting satellite clock bias based on robust fitting of ARMA models. GPS Solutions 26:3. https://doi.org/10.1007/s10291-021-01182-3
Zhao L, Li N, Li H, Wang RL, Li MH (2021) BDS satellite clock prediction considering periodic variations. Remote Sensing 13(20):4058. https://doi.org/10.3390/rs13204058
Zheng ZY, Dang YM, Lu XS, Xu WM (2010) Prediction model with periodic item and its application to the prediction of GPS satellite clock bias. Acta Astronomica Sinica 51(1):95-102. https://doi.org/10.15940/j.cnki.0001-5245.2010.01.012
Zheng ZY, Lu XS, Chen YQ (2008) Improved grey model and application in Real-Time GPS satellite clock bias prediction. In: 2008 Fourth International Conference on Natural Computation, pp. 419-423. https://doi.org/10.1109/ICNC.2008.630
Zhu XW, Xiao H, Yong SW, Zhuang ZW (2008) The Kalman algorithm used for satellite clock offset prediction and its performance analysis. Journal of Astronautics 29(3):966-970.